行政院國家科學委員會專題研究計畫　期中進度報告

多元指數一般型 II 設限資料的完全統計推論與預測(1/2)

計畫類別： 個別型計畫
計畫編號： NSC92-2118-M-032-001-
執行期間： 92 年 08 月 01 日至 93 年 07 月 31 日
執行單位： 淡江大學數學系

計畫主持人： 林千代

報告類型： 精簡報告

處理方式： 本計畫可公開查詢

中　華　民　國 93 年 5 月 12 日

中文摘要:

當存活資料或工業統計失敗資料是具有一元單參數或雙參數指數分配的一般型 II 設限資料時，文獻上並沒有適當的關鍵統計量的完全分配結果存在以供我們取得參數的完全區間估計或缺失值的預測區間．同樣的窘態也發生在多元單參數或雙參數指數分配的一般型 II 設限資料時. 過去四十年，學者多以最大概似估計法處理這方面資料，但是往往須要特別程式解決，因而很多建立參數信賴區間及假設檢定研究問題多無法發展，結果多以逼近法來彌補．有鑑於此缺失,我們利用 Huffer and Lin (2001) 針對一般化留間隔線性組合的機率問題所提出的演算法，針對雙參數指數分配的雙邊型 II 設限樣本之最佳線性不偏估計式所形成的關鍵統計量計算出其實際的百分比數位置，進而建造出參數的完全區間區間或缺失值的預測區間．我們會將所得的結果和以最大概似估計式逼近法所建立區間區間或缺失值的預測區間相互比較．最後我們會呈現以所提出的推論計算方法所應用的實例．

計畫成果自評部分:

本年度所作研究已經獲得學術期刊接受.
N. Balakrishnan, C. T. Lin, and P. S. Chan. (2004). Exact Inference and Prediction for K-Sample Two-Parameter Exponential Case Under General Type-II Censoring. Journal of Statistical Computation and Simulation. (in press) (SCI) NSC-92-2118-M-032-001.
相信後續研究應該可以在學術期刊發表.

Exact Inference and Prediction for K-Sample Two-Parameter Exponential Case Under doubly Type-II Censoring

Chien-Tai Lin, Tamkang University, Taiwan

Abstract

Exact inference for the location and scale parameters as well as prediction intervals for the K-sample exponential case under doubly Type-II censored samples are derived using an algorithm of Huffer and Lin (2001). This approach provides a simple way to determine the exact percentage points of the pivotal quantities based on the best linear unbiased estimators in order to develop exact inference for the location and scale parameters as well as to construct exact prediction intervals for failure times unobserved in the i-th sample. Similarly, exact prediction intervals for failure times of units from a future sample can also be easily obtained. A comparison is then made with the approximate inference based on the maximum likelihood estimators. Finally, we present an example to illustrate all the methods of inference developed in this report.

Keywords

Best linear unbiased estimator; exponential distribution; K-sample doubly Type-II censored sample; normalized spacings.

1 Introduction

Exponential distribution is a commonly used model in life-testing and reliability studies. A very wide variety of properties and applications of the exponential distribution can be found in the books by Nelson (1982), Lawless (1982), Bain and Engelhardt (1991), Johnson, Kotz and Balakrishnan (1994), Balakrishnan and Basu (1995), and Meeker and Escobar (1998). It is well known that the normalized spacings from an exponential distribution are independent and identically distributed as exponential; see, for example, David (1981) and Arnold, Balakrishnan and Nagaraja (1992). This property, for example, allows the development of exact chi-square confidence intervals for the scale parameter and exact F confidence intervals for the location parameter of a two-parameter exponential distribution based on Type-II right censored samples. It has also been utilized by many authors to construct exact prediction intervals for failure times of items having observed the first $n-s$ failures from a sample of n items placed on a life test; see, for example, Lawless (1971, 1977), Likeš (1974) and Lingappaiah (1973). It should be mentioned here that a general survey of prediction problems has been given by Patel (1989), Kaminsky and Nelson (1998) and Nagaraja (1995), with the latter presenting a comprehensive review of results on prediction problems for exponential distribution wherein Box's (1954) results on linear functions of chi-square random variables is extensively used. However, when the available sample from the life-testing experiment is doubly Type-II censored, no exact distributional result is available for the appropriate pivotal quantity useful for this prediction in the two-parameter exponential distribution. For this reason, Lin and Balakrishnan (2003) recently used Huffer and Lin's (2001) algorithm for this specific problem.

In this report, we develop exact confidence intervals for the exponential location and scale parameters based on K-sample doubly Type-II censored exponential data by utilizing an algorithm of Huffer and Lin (2001). We also construct exact prediction intervals for failure times unobserved in the i-th sample as well as for failure times from a future sample. We then make a comparison with the approximate inference based on the maximum likelihood estimators. Finally, we illustrate all the methods of inference developed here with a data from Nelson (1982, Ch. 10, Table 4.1).

2 Exact Inference and Prediction based on BLUEs

We suppose that K independent doubly Type-II censored samples are available from an exponential distribution with probability density function $f(y;\mu,\sigma) = \frac{1}{\sigma}e^{-(y-\mu)/\sigma}$, $y \geq \mu$, $\sigma > 0$. Denote the total number of observations in the i-th sample by n_i, and the vector of order statistics observed from the i-th sample by $\boldsymbol{Y}_i = (Y_{i(r_i+1:n_i)}, Y_{i(r_i+2:n_i)}, \ldots, Y_{i(n_i-s_i:n_i)})^T$, $1 \leq r_i+1 \leq n_i - s_i \leq n_i$, $i = 1, 2, \ldots, K$; that is, r_i smallest and s_i largest order statistics have been censored in the i-th sample. Then, the vector $\boldsymbol{Y} = (\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_K)^T$ has its mean vector as $\boldsymbol{\mu} = E(\boldsymbol{Y}) = (\mu\mathbf{1}+\sigma\boldsymbol{\alpha}_1, \mu\mathbf{1}+\sigma\boldsymbol{\alpha}_2, \ldots, \mu\mathbf{1}+\sigma\boldsymbol{\alpha}_K)^T$ and variance-covariance matrix as $Var(\boldsymbol{Y}) = \sigma^2\boldsymbol{\Sigma} = \sigma^2\ \mathrm{Diag}(\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_K)$, where $\boldsymbol{\alpha}_i = (\mu_{i(r_i+1:n_i)}, \ldots, \mu_{i(n_i-s_i:n_i)})^T$ with $\mu_{i(\ell:n_i)} = \sum_{j=1}^{\ell} 1/(n_i-j+1)$, $r_i+1 \leq \ell \leq n_i - s_i$, and

$$\boldsymbol{\Sigma}_i = \begin{pmatrix} \sigma_{i(r_i+1,r_i+1:n_i)} & \sigma_{i(r_i+1,r_i+2:n_i)} & \cdots & \sigma_{i(r_i+1,n_i-s_i:n_i)} \\ \sigma_{i(r_i+1,r_i+2:n_i)} & \sigma_{i(r_i+2,r_i+2:n_i)} & \cdots & \sigma_{i(r_i+2,n_i-s_i:n_i)} \\ \vdots & \vdots & \cdots & \vdots \\ \sigma_{i(r_i+1,n_i-s_i:n_i)} & \sigma_{i(r_i+2,n_i-s_i:n_i)} & \cdots & \sigma_{i(n_i-s_i,n_i-s_i:n_i)} \end{pmatrix}$$

with $\sigma_{i(\ell,\ell:n_i)} = \sigma_{i(\ell,q:n_i)} = \sum_{j=1}^{\ell} 1/(n_i-j+1)^2$, $r_i+1 \leq \ell \leq q \leq n_i - s_i$; see, for example, David (1981) and Arnold, Balakrishnan and Nagaraja (1992).

We minimize the generalized variance $W = (\boldsymbol{Y}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{\mu})$ to obtain the BLUEs of μ and σ as

$$\mu^* = \frac{\left(\sum_{i=1}^K \boldsymbol{\alpha}_i^T\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\alpha}_i\right)\left(\sum_{i=1}^K \mathbf{1}^T\boldsymbol{\Sigma}_i^{-1}\boldsymbol{Y}_i\right) - \left(\sum_{i=1}^K \mathbf{1}^T\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\alpha}_i\right)\left(\sum_{i=1}^K \boldsymbol{\alpha}_i^T\boldsymbol{\Sigma}_i^{-1}\boldsymbol{Y}_i\right)}{\left(\sum_{i=1}^K \mathbf{1}^T\boldsymbol{\Sigma}_i^{-1}\mathbf{1}\right)\left(\sum_{i=1}^K \boldsymbol{\alpha}_i^T\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\alpha}_i\right) - \left(\sum_{i=1}^K \mathbf{1}^T\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\alpha}_i\right)^2} \tag{1}$$

and

$$\sigma^* = \frac{\left(\sum_{i=1}^K \mathbf{1}^T\boldsymbol{\Sigma}_i^{-1}\mathbf{1}\right)\left(\sum_{i=1}^K \boldsymbol{\alpha}_i^T\boldsymbol{\Sigma}_i^{-1}\boldsymbol{Y}_i\right) - \left(\sum_{i=1}^K \mathbf{1}^T\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\alpha}_i\right)\left(\sum_{i=1}^K \mathbf{1}^T\boldsymbol{\Sigma}_i^{-1}\boldsymbol{Y}_i\right)}{\left(\sum_{i=1}^K \mathbf{1}^T\boldsymbol{\Sigma}_i^{-1}\mathbf{1}\right)\left(\sum_{i=1}^K \boldsymbol{\alpha}_i^T\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\alpha}_i\right) - \left(\sum_{i=1}^K \mathbf{1}^T\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\alpha}_i\right)^2}. \tag{2}$$

Let us denote $N = \sum_{i=1}^K n_i$, $R = \sum_{i=1}^K r_i$, $S = \sum_{i=1}^K s_i$, $Q = \sum_{i=1}^K q_i$, $T = \sum_{i=1}^K t_i$ and $U = \sum_{i=1}^K u_i$, where $q_i = \left(\sum_{\ell=n_i-r_i}^{n_i} \frac{1}{\ell}\right)^2 / \left(\sum_{\ell=n_i-r_i}^{n_i} \frac{1}{\ell^2}\right)$, $t_i = \left(\sum_{\ell=n_i-r_i}^{n_i} \frac{1}{\ell}\right) / \left(\sum_{\ell=n_i-r_i}^{n_i} \frac{1}{\ell^2}\right)$, and $u_i = 1/\left(\sum_{\ell=n_i-r_i}^{n_i} \frac{1}{\ell^2}\right)$ for $i = 1, 2, \ldots, K$. Observe that $t_i^2 = q_i u_i$ for $i = 1, 2, \ldots, K$. Since $\boldsymbol{\alpha}_i^T \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\alpha}_i = n_i - r_i - s_i - 1 + q_i$, $\mathbf{1}^T \boldsymbol{\Sigma}_i^{-1} \mathbf{1} = u_i$, and $\mathbf{1}^T \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\alpha}_i = t_i$, we readily have $A_1 = \sum_{i=1}^K \boldsymbol{\alpha}_i^T \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\alpha}_i = N - R - S - K + Q$, $A_2 = \sum_{i=1}^K \mathbf{1}^T \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\alpha}_i = T$, $A_3 = \sum_{i=1}^K \mathbf{1}^T \boldsymbol{\Sigma}_i^{-1} \mathbf{1} = U$, and $D = A_1 A_3 - A_2^2$. Note that D is positive since $T^2 = \left(\sum_{i=1}^K t_i\right)^2 = \left(\sqrt{q_i}\sqrt{u_i}\right)^2 \leq QU$ by Cauchy-Schwarz inequality. Also, we find that

$$
\begin{aligned}
(\mathbf{1}^T \boldsymbol{\Sigma}_i^{-1})_j &= \begin{cases} \frac{1}{\sum_{\ell=n_i-r_i}^{n_i} \frac{1}{\ell^2}} & \text{for } j = r_i + 1, \\ 0 & \text{for } r_i + 2 \leq j \leq n_i - s_i, \end{cases} \\
&= \begin{cases} u_i & \text{for } j = r_i + 1, \\ 0 & \text{for } r_i + 2 \leq j \leq n_i - s_i, \end{cases}
\end{aligned}
$$

and

$$
\begin{aligned}
(\boldsymbol{\alpha}_i^T \boldsymbol{\Sigma}_i^{-1})_j &= \begin{cases} \frac{\sum_{\ell=n_i-r_i}^{n_i} \frac{1}{\ell}}{\sum_{\ell=n_i-r_i}^{n_i} \frac{1}{\ell^2}} - (n_i - r_i - 1) & \text{for } j = r_i + 1, \\ 1 & \text{for } r_i + 2 \leq j \leq n_i - s_i - 1, \\ s_i + 1 & \text{for } j = n_i - s_i, \end{cases} \\
&= \begin{cases} t_i - (n_i - r_i - 1) & \text{for } j = r_i + 1, \\ 1 & \text{for } r_i + 2 \leq j \leq n_i - s_i - 1, \\ s_i + 1 & \text{for } j = n_i - s_i, \end{cases}
\end{aligned}
$$

which readily yield

$$
\sum_{i=1}^K \mathbf{1}^T \boldsymbol{\Sigma}_i^{-1} \boldsymbol{Y}_i = u_1 Y_{1(r_1+1:n_1)} + u_2 Y_{2(r_2+1:n_2)} + \ldots + u_K Y_{K(r_K+1:n_K)}
$$

and

$$
\sum_{i=1}^K \boldsymbol{\alpha}_i^T \boldsymbol{\Sigma}_i^{-1} \boldsymbol{Y}_i = \sum_{i=1}^K \left\{ [t_i - (n_i - r_i - 1)] Y_{i(r_i+1:n_i)} + \sum_{j=r_i+2}^{n_i-s_i-1} Y_{i(j:n_i)} + (s_i + 1) Y_{i(n_i-s_i:n_i)} \right\}.
$$

Substituting all these expressions into (1) and (2), we obtain the BLUEs μ^* and σ^* as

$$
\begin{aligned}
\mu^* &= \frac{(N - R - S - K + Q) \sum_{i=1}^K u_i Y_{i(r_i+1:n_i)}}{D} \\
&\quad - \frac{T \sum_{i=1}^K \left\{ [t_i - (n_i - r_i - 1)] Y_{i(r_i+1:n_i)} + \sum_{j=r_i+2}^{n_i-s_i-1} Y_{i(j:n_i)} + (s_i + 1) Y_{i(n_i-s_i:n_i)} \right\}}{D} \\
&= \frac{\sum_{i=1}^K [(N - R - S - K + Q) u_i - T t_i] Y_{i(r_i+1:n_i)}}{D} \\
&\quad - \frac{T \sum_{i=1}^K [\sum_{j=r_i+2}^{n_i-s_i-1} Y_{i(j:n_i)} + (s_i + 1) Y_{i(n_i-s_i:n_i)} - (n_i - r_i - 1) Y_{i(r_i+1:n_i)}]}{D}
\end{aligned} \tag{3}
$$

and

$$\sigma^* = \frac{U\sum_{i=1}^{K}\left\{[t_i-(n_i-r_i-1)]Y_{i(r_i+1:n_i)}+\sum_{j=r_i+2}^{n_i-s_i-1}Y_{i(j:n_i)}+(s_i+1)Y_{i(n_i-s_i:n_i)}\right\}}{D} - \frac{T\sum_{i=1}^{K}u_iY_{i(r_i+1:n_i)}}{D}. \tag{4}$$

Denote $S_{i(j)} = (n_i - j + 1)\left(Y_{i(j:n_i)} - Y_{i(j-1:n_i)}\right)$ for $j = 1, 2, \ldots, n_i - s_i$ and $i = 1, 2, \ldots, K$, with the convention that $Y_{i(0:n_i)} = \mu$. By the spacings property mentioned earlier in Section 1, we readily have these $S_{i(j)}$'s to be i.i.d. exponential (with scale parameter σ) random variables. Then, the BLUEs of μ and σ in (3) and (4) can be simplified and rewritten as

$$\begin{aligned}\mu^* &= \frac{\sum_{i=1}^{K}[(N-R-S-K+Q)u_i-Tt_i](\mu+\sum_{j=1}^{r_i+1}\frac{1}{n_i-j+1}S_{i(j)})}{D} \\ &\quad -\frac{T\sum_{i=1}^{K}\sum_{j=r_i+2}^{n_i-s_i}S_{i(j)}}{D} \\ &= \mu+\frac{\sum_{i=1}^{K}\sum_{j=1}^{r_i+1}\frac{(N-R-S-K+Q)u_i-Tt_i}{n_i-j+1}S_{i(j)}-T\sum_{i=1}^{K}\sum_{j=r_i+2}^{n_i-s_i}S_{i(j)}}{D}\end{aligned}$$

and

$$\begin{aligned}\sigma^* &= \frac{U\sum_{i=1}^{K}\sum_{j=r_i+2}^{n_i-s_i}S_{i(j)}-\sum_{i=1}^{K}(Tu_i-Ut_i)Y_{i(r_i+1:n_i)}}{D} \\ &= \frac{U\sum_{i=1}^{K}\sum_{j=r_i+2}^{n_i-s_i}S_{i(j)}-\sum_{i=1}^{K}(Tu_i-Ut_i)\left(\mu+\sum_{j=1}^{r_i+1}\frac{S_{i(j)}}{n_i-j+1}\right)}{D} \\ &= \frac{U\sum_{i=1}^{K}\sum_{j=r_i+2}^{n_i-s_i}S_{i(j)}-\sum_{i=1}^{K}\sum_{j=1}^{r_i+1}\frac{Tu_i-Ut_i}{n_i-j+1}S_{i(j)}}{D}.\end{aligned}$$

The variances and covariance of the BLUEs μ^* and σ^* are obtained as

$$\begin{aligned}Var(\mu^*) &= \frac{(N-R-S-K+Q)\sigma^2}{D}, \\ Var(\sigma^*) &= \frac{\sigma^2}{D^2}\sum_{i=1}^{K}\left\{\sum_{j=1}^{r_i+1}\frac{(-Tu_i+Ut_i)^2}{(n_i-j+1)^2}+\sum_{j=r_i+2}^{n_i-s_i}U^2\right\}=\frac{U\sigma^2}{D},\end{aligned}$$

and

$$Cov(\mu^*,\sigma^*) = \frac{-T\sigma^2}{D}.$$

Using the property of normalized spacings and independence of the samples, we can apply the algorithm outlined in Section 2 to determine the exact value of d satisfying

$$P\left(\frac{\sigma^*}{\sigma}>d\right)=P\left(\frac{U\sum_{i=1}^{K}\sum_{j=r_i+2}^{n_i-s_i}Z_{i(j)}-\sum_{i=1}^{K}\sum_{j=1}^{r_i+1}\frac{Tu_i-Ut_i}{n_i-j+1}Z_{i(j)}}{D}>d\right)=\alpha, \tag{5}$$

and

$$
\begin{aligned}
& P\left(\frac{\mu^* - \mu}{\sigma^*} > d\right) \\
= \; & P\left(\frac{\sum_{i=1}^{K}\sum_{j=1}^{r_i+1}\frac{(N-R-S-K+Q)u_i - Tt_i}{n_i-j+1}Z_{i(j)} - T\sum_{i=1}^{K}\sum_{j=r_i+2}^{n_i-s_i}Z_{i(j)}}{U\sum_{i=1}^{K}\sum_{j=r_i+2}^{n_i-s_i}Z_{i(j)} - \sum_{i=1}^{K}\sum_{j=1}^{r_i+1}\frac{Tu_i-Ut_i}{n_i-j+1}Z_{i(j)}} > d\right) \\
= \; & P\left(\sum_{i=1}^{K}\sum_{j=1}^{r_i+1}\frac{(N-R-S-K+Q+Td)u_i - (T+Ud)t_i}{n_i-j+1}Z_{i(j)}\right. \\
& \qquad \left. -\sum_{i=1}^{K}\sum_{j=r_i+2}^{n_i-s_i}(T+Ud)Z_{i(j)} > 0\right) = \alpha \qquad (6)
\end{aligned}
$$

for any $0 < \alpha < 1$, where $Z_{i(j)}$'s are i.i.d. standard exponential random variables for $i = 1, \ldots, K$ and $j = 1, \ldots, n_i - s_i$.

In a similar manner, we can construct exact prediction intervals for a failure time unobserved in the i-th sample, viz. $Y_{i(h:n_i)}$ for $n_i - s_i < h \leq n_i$, $i = 1, \ldots, K$, by finding the exact value of d such that

$$
\begin{aligned}
\alpha &= P\left(\frac{Y_{i(h:n_i)} - Y_{i(n_i-s_i:n_i)}}{\sigma^*} > d\right) \\
&= P\left(\sum_{j=n_i-s_i+1}^{h}\frac{1}{n_i-j+1}Z_{i(j)}\right. \\
&\qquad \left. -\frac{d}{D}\left[U\sum_{i=1}^{K}\sum_{j=r_i+2}^{n_i-s_i}Z_{i(j)} - \sum_{i=1}^{K}\sum_{j=1}^{r_i+1}\frac{Tu_i-Ut_i}{n_i-j+1}Z_{i(j)}\right] > 0\right). \qquad (7)
\end{aligned}
$$

Thus, for a specified α, we can determine values of d_1 and d_2 such that

$$
P\left(\frac{Y_{i(h:n_i)} - Y_{i(n_i-s_i:n_i)}}{\sigma^*} > d_1\right) = \frac{\alpha}{2} \text{ and } P\left(\frac{Y_{i(h:n_i)} - Y_{i(n_i-s_i:n_i)}}{\sigma^*} > d_2\right) = 1 - \frac{\alpha}{2};
$$

then, an exact $100(1-\alpha)\%$ prediction interval for the unobserved failure time $Y_{i(h:n_i)}$ is given by $(Y_{i(n_i-s_i:n_i)} + d_2\sigma^*, Y_{i(n_i-s_i:n_i)} + d_1\sigma^*)$.

Proceeding in an analogous manner, we can also construct exact prediction intervals for failure times from a future sample, viz. $Y_{K+1(h:n_{K+1})}$, $h = 1, \ldots, n_{K+1}$, by finding the exact value of d such that

$$
\begin{aligned}
\alpha &= P\left(\frac{Y_{K+1(h:n_{K+1})} - \mu^*}{\sigma^*} > d\right) \\
&= P\left(\sum_{j=1}^{h}\frac{1}{n_{K+1}-j+1}Z_{K+1(j)}\right. \\
&\quad \left. - \frac{\sum_{i=1}^{K}\sum_{j=1}^{r_i+1}\frac{(N-R-S-K+Q-dT)u_i-(T-dU)t_i}{n_i-j+1}Z_{i(j)} - (T-dU)\sum_{i=1}^{K}\sum_{j=r_i+2}^{n_i-s_i}Z_{i(j)}}{D} > 0\right).
\end{aligned}
\qquad (8)
$$

Thus, an exact $100(1-\alpha)\%$ prediction interval for the failure time $Y_{K+1(h:n_{K+1})}$ from a future sample is given by $(\mu^* + d_4\sigma^*, \mu^* + d_3\sigma^*)$, where

$$P\left(\frac{Y_{K+1(h:n_{K+1})} - \mu^*}{\sigma^*} > d_3\right) = \frac{\alpha}{2} \quad \text{and} \quad P\left(\frac{Y_{K+1(h:n_{K+1})} - \mu^*}{\sigma^*} > d_4\right) = 1 - \frac{\alpha}{2}.$$

3 Maximum Likelihood Estimation

The likelihood function of μ and σ is

$$\begin{aligned} L(\mu,\sigma) &= \prod_{i=1}^{K} \frac{n_i!}{r_i! s_i!} \left[1 - \exp\left(-x_{i(r_i+1:n_i)}\right)\right]^{r_i} \left[\exp\left(-x_{i(n_i-s_i:n_i)}\right)\right]^{s_i} \\ &\quad \times \frac{1}{\sigma^{n_i-r_i-s_i}} \exp\left[-\sum_{j=r_i+1}^{n_i-s_i} x_{i(j:n_i)}\right], \quad \mu \le \min_{1\le i\le K} y_{i(r_i+1:n_i)},\ \sigma > 0, \quad (9)\end{aligned}$$

where $x_{i(j:n_i)} = (y_{i(j:n_i)} - \mu)/\sigma, j = r_i+1, \ldots, n_i - s_i, i = 1, 2, \ldots, K$. The derivatives of the log-likelihood with respect to the parameters are obtained from (9) to be

$$\frac{\partial \log L}{\partial \mu} = -\frac{1}{\sigma}\left\{\sum_{i=1}^{K} r_i \frac{\exp\left(-x_{i(r_i+1:n_i)}\right)}{1-\exp\left(-x_{i(r_i+1:n_i)}\right)} - (N-R)\right\} \quad (10)$$

and

$$\begin{aligned} \frac{\partial \log L}{\partial \sigma} &= -\frac{1}{\sigma}\Bigg\{(N-R-S) + \sum_{i=1}^{K} r_i x_{i(r_i+1:n_i)} \frac{\exp\left(-x_{i(r_i+1:n_i)}\right)}{1-\exp\left(-x_{i(r_i+1:n_i)}\right)} \\ &\qquad - \sum_{i=1}^{K} s_i x_{i(n_i-s_i:n_i)} - \sum_{i=1}^{K}\sum_{j=r_i+1}^{n_i-s_i} x_{i(j:n_i)}\Bigg\}, \quad (11)\end{aligned}$$

respectively. The second derivatives of the log-likelihood function are similarly obtained as

$$\begin{aligned} \frac{\partial^2 \log L}{\partial \mu^2} &= -\frac{1}{\sigma^2}\left\{\sum_{i=1}^{K} r_i \frac{\exp\left(-x_{i(r_i+1:n_i)}\right)}{1-\exp\left(-x_{i(r_i+1:n_i)}\right)} + \sum_{i=1}^{K} r_i \left(\frac{\exp\left(-x_{i(r_i+1:n_i)}\right)}{1-\exp\left(-x_{i(r_i+1:n_i)}\right)}\right)^2\right\} \\ &= -\frac{1}{\sigma^2}\sum_{i=1}^{K} r_i \frac{\exp\left(-x_{i(r_i+1:n_i)}\right)}{\left[1-\exp\left(-x_{i(r_i+1:n_i)}\right)\right]^2}, \quad (12)\end{aligned}$$

$$\frac{\partial^2 \log L}{\partial \mu \partial \sigma} = -\frac{1}{\sigma^2}\left\{\sum_{i=1}^{K} r_i x_{i(r_i+1:n_i)} \frac{\exp\left(-x_{i(r_i+1:n_i)}\right)}{\left[1-\exp\left(-x_{i(r_i+1:n_i)}\right)\right]^2}\right\}, \quad (13)$$

and

$$\frac{\partial^2 \log L}{\partial \sigma^2} = \frac{1}{\sigma^2}\Bigg\{(N-R-S) + \sum_{i=1}^{K} r_i x_{i(r_i+1:n_i)} \frac{\exp\left(-x_{i(r_i+1:n_i)}\right)}{1-\exp\left(-x_{i(r_i+1:n_i)}\right)}$$

$$
\begin{aligned}
& -\sum_{i=1}^{K} s_i x_{i(n_i-s_i:n_i)} - \sum_{i=1}^{K}\sum_{j=r_i+1}^{n_i-s_i} x_{i(j:n_i)}\Bigg\} \\
& +\frac{1}{\sigma^2}\Bigg\{\sum_{i=1}^{K} r_i x_{i(r_i+1:n_i)}\left(\frac{\exp\left(-x_{i(r_i+1:n_i)}\right)}{1-\exp\left(-x_{i(r_i+1:n_i)}\right)} - x_{i(r_i+1:n_i)}\right. \\
& \left.\times\left[\frac{\exp\left(-x_{i(r_i+1:n_i)}\right)}{1-\exp\left(-x_{i(r_i+1:n_i)}\right)} + \left(\frac{\exp\left(-x_{i(r_i+1:n_i)}\right)}{1-\exp\left(-x_{i(r_i+1:n_i)}\right)}\right)^2\right]\right) \\
& -\sum_{i=1}^{K} s_i x_{i(n_i-s_i:n_i)} - \sum_{i=1}^{K}\sum_{j=r_i+1}^{n_i-s_i} x_{i(j:n_i)}\Bigg\} \\
= \; & -\frac{1}{\sigma^2}\Bigg\{\sum_{i=1}^{K} r_i x^2_{i(r_i+1:n_i)}\frac{\exp\left(-x_{i(r_i+1:n_i)}\right)}{\left[1-\exp\left(-x_{i(r_i+1:n_i)}\right)\right]^2} + (N-R-S)\Bigg\}. \qquad (14)
\end{aligned}
$$

The maximum likelihood estimates of μ and σ can be obtained from (10) and (11) using Newton-Raphson method that requires the computation of second derivatives in (12), (13) and (14) at every iteration.

4 Numerical Illustration

Consider Nelson's data (1982, Ch. 10, Table 4.1) which give 60 times to breakdown in minutes of an insulating fluid subjected to high voltage stress. The failure times were observed in the form of six groups each with ten insultating fluids. For the purpose of illustrating the methods of inference detailed in the previous sections, we consider the following six doubly Type-II censored samples with $(n_i, r_i, s_i), i = 1, \ldots, 6$, taken respectively to be $(10,2,1), (10,1,1), (10,1,1), (10,1,1), (10,1,2), (10,1,2)$:

Group 1	–	–	1.54	1.70	1.82	1.89	2.17	2.24	4.03	–
Group 2	–	0.18	0.55	0.66	0.71	1.30	1.63	2.17	2.75	–
Group 3	–	0.64	0.82	0.93	1.08	1.99	2.06	2.15	2.57	–
Group 4	–	0.06	0.50	0.70	1.17	2.80	3.57	3.72	3.82	–
Group 5	–	0.78	0.80	1.08	1.13	2.44	3.17	5.55	–	–
Group 6	–	1.49	1.56	2.10	2.12	3.83	3.97	5.13	–	–

For the above data, we find the maximum likelihood estimates of μ and σ from (10) and (11) to be $\hat{\mu} = -0.016$ and $\hat{\sigma} = 2.449$, respectively. Furthermore, through a Monte Carlo simulation study, the variance-covariance matrix of these MLEs is determined to be

$$
Var\begin{pmatrix}\hat{\mu}\\ \hat{\sigma}\end{pmatrix} = \hat{\sigma}^2\begin{pmatrix}0.0081 & -0.00737\\ -0.00737 & 0.02536\end{pmatrix} = \begin{pmatrix}0.04858 & -0.04420\\ -0.04420 & 0.15210\end{pmatrix}.
$$

Now, upon using the asymptotic normality of the MLEs, we find an approximate 95% confidence interval for μ (based on MLEs) to be

$$
\left[\hat{\mu} - 1.96\sqrt{0.04858}, \hat{\mu} + 1.96\sqrt{0.04858}\right] = [-0.416, 0.385],
$$

and an approximate 95% confidence interval for σ (based on MLEs) to be

$$\left[\hat{\sigma} - 1.96\sqrt{0.15210}, \hat{\sigma} + 1.96\sqrt{0.15210}\right] = [1.685, 3.213].$$

Next, we obtain the BLUEs of μ and σ as $\mu^* = 0.23812$ and $\sigma^* = 2.17461$ from (3) and (4). Applying the algorithm of Huffer and Lin (2001) we find

$$P\left(\frac{\sigma^*}{\sigma} > 0.71215\right) = 0.975 \quad \text{and} \quad P\left(\frac{\sigma^*}{\sigma} > 1.33555\right) = 0.025$$

using which we obtain an exact 95% confidence interval for σ to be

$$\left[\frac{\sigma^*}{1.33555}, \frac{\sigma^*}{0.71215}\right] = [1.62825, 3.05358].$$

We also find

$$P\left(\frac{\mu^* - \mu}{\sigma^*} > -0.11249\right) = 0.975 \quad \text{and} \quad P\left(\frac{\mu^* - \mu}{\sigma^*} > 0.18210\right) = 0.025$$

using which we obtain an exact 95% confidence interval for μ to be

$$[0.23812 - 0.18210 \times 2.17461, 0.23812 + 0.11249 \times 2.17461] = [-0.15788, 0.48274].$$

It is important to mention that the confidence intervals based on BLUEs are exact while those based on MLEs are approximate. Yet, they are close numerically; furthermore, both confidence intervals for μ include 0, indicating that one may as well use a one-parameter exponential for the data at hand.

Similarly, with $\boldsymbol{A}$ in (7), we find

$$P\left(\frac{Y_{5(9)} - Y_{5(8)}}{\sigma^*} > 0.01266\right) = 0.975 \quad \text{and} \quad P\left(\frac{Y_{5(9)} - Y_{5(8)}}{\sigma^*} > 1.93366\right) = 0.025$$

and

$$P\left(\frac{Y_{5(10)} - Y_{5(8)}}{\sigma^*} > 0.17051\right) = 0.975 \quad \text{and} \quad P\left(\frac{Y_{5(10)} - Y_{5(8)}}{\sigma^*} > 4.62437\right) = 0.025.$$

Using these results, the exact 95% prediction intervals for the last two failures in Group 5, viz. $Y_{5(9)}$ and $Y_{5(10)}$, are obtained as

$$[5.55 + 2.17461 \times 0.01266,\ 5.55 + 2.17461 \times 1.93366] = [5.57753,\ 9.75496],$$

and

$$[5.55 + 2.17461 \times 0.17051,\ 5.55 + 2.17461 \times 4.62437] = [5.92079,\ 15.60620],$$

respectively.

Suppose a sample of 3 insulating fluids is to be tested in future and that we wish to predict the corresponding failure times. Then, with $n_7 = 3$, we find

$$P\left(\frac{Y_{7(1)} - \mu^*}{\sigma^*} > -0.06588\right) = 0.975 \quad \text{and} \quad P\left(\frac{Y_{7(1)} - \mu^*}{\sigma^*} > 1.27539\right) = 0.025,$$

$$P\left(\frac{Y_{7(2)} - \mu^*}{\sigma^*} > 0.07978\right) = 0.975 \quad \text{and} \quad P\left(\frac{Y_{7(2)} - \mu^*}{\sigma^*} > 2.47882\right) = 0.025,$$

and

$$P\left(\frac{Y_{7(3)} - \mu^*}{\sigma^*} > 0.33754\right) = 0.975 \quad \text{and} \quad P\left(\frac{Y_{7(3)} - \mu^*}{\sigma^*} > 5.04952\right) = 0.025$$

with $\boldsymbol{A}$ in (8). Making use of all these results, we obtain the exact 95% prediction intervals for the ordered failure times $Y_{7(1)}, Y_{7(2)}, Y_{7(3)}$ from a future sample of size $n_7 = 3$ to be, respectively,

$$[0.23812 - 2.17461 \times 0.06588,\ 0.23812 + 2.17461 \times 1.27539] = [0.09486,\ 3.01160]\,,$$

$$[0.23812 + 2.17461 \times 0.07978,\ 0.23812 + 2.17461 \times 2.47882] = [0.41161,\ 5.62859]\,,$$

and

$$[0.23812 + 2.17461 \times 0.33754,\ 0.23812 + 2.17461 \times 5.04952] = [0.97214,\ 11.21886]\,.$$

References

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1992). *A First Course in Order Statistics*, New York: John Wiley & Sons.

Bain, L. J. and Engelhardt, M. (1991). *Statistical Analysis of Reliability and Life-Testing Models*, Second edition, New York: Marcel Dekker.

Balakrishnan, N. and Basu, A. P. (Eds.) (1995). *The Exponential Distribution: Theory, Methods and Applications*, Langhorne, PA: Gordon and Breach Science Publishers.

Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification, *Annals of Mathematical Statistics*, **25**, 290–302.

David, H. A. (1981). *Order Statistics*, Second edition, New York: John Wiley & Sons.

Huffer, F. W. (1988). Divided differences and the joint distribution of linear combinations of spacings, *Journal of Applied Probability*, **25**, 346–354.

Huffer, F. W. and Lin, C. T. (1999). An approach to computations involving spacings with applications to the scan statistic, In *Scan Statistics and Applications* (Eds., J. Glaz and N. Balakrishnan), pp. 141–163, Boston: Birkhäuser.

Huffer, F. W. and Lin, C. T. (2001). Computing the joint distribution of general linear combinations of spacings or exponential variates, *Statistica Sinica*, **11**, 1141–1157.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions – Vol. 1*, Second edition, New York: John Wiley & Sons.

Kaminsky, K. S. and Nelson, P. I. (1998). Prediction of order statistics, In *Handbook of Statistics – 17: Order Statistics: Applications* (Eds., N. Balakrishnan and C. R. Rao), pp. 431–450, Amsterdam: North-Holland.

Lawless, J. F. (1971). A prediction problem concerning samples from the exponential distribution, with application in life testing, *Technometrics*, **13**, 725–730.

Lawless, J. F. (1977). Prediction intervals for the two-parameter exponential distribution, *Technometrics*, **17**, 255–261.

Lawless, J. F. (1982). *Statistical Models & Methods for Lifetime Data*, New York: John Wiley & Sons.

Likeš, J. (1974). Prediction of sth ordered observation for the two-parameter exponential distribution, *Technometrics*, **16**, 241–244.

Lin, C. T. (1993). The computation of probabilities which involve spacings, with applications to the scan statistic, *Ph.D. Dissertation*, Department of Statistics, Florida State University, Tallahassee, FL.

Lin, C. T. and N. Balakrishnan. (2003). Exact prediction intervals for exponential distributions based on doubly Type-II censored samples. *Journal of Applied Statistics* **30**, 783–801.

Lingappaiah, G. S. (1973). Prediction in exponential life testing, *Canadian Journal of Statistics*, **1**, 113–118.

Meeker, W. Q. and Escobar, L. A. (1998). *Statistical Methods for Reliability Data*, New York: John Wiley & Sons.

Nagaraja, H. N. (1995). Prediction problems, In *The Exponential Distribution: Theory, Methods and Applications* (Eds., N. Balakrishnan and A. P. Basu), pp. 139–163, Langhorne, PA: Gordon and Breach Science Publishers.

Nelson, W. (1982). *Applied Life Data Analysis*, New York: John Wiley & Sons.

Patel, J. K. (1989). Prediction intervals – A review, *Communications in Statistics — Theory and Methods*, **18**, 2393–2465.